Appendix/Electronic Supplement/Download

Appendix 1: Methods for comparisons of two-point *vs*. multiple-point fluxes

All fluxes were estimated using *flux* (*Jurasinski* et al., 2014) in R. Therefore, they were all estimated by fitting a linear regression model (no non-linear models). MP-fluxes were calculated with the possibility to exclude one concentration measurement (outlier detection), because this option is one important advantage of MP-measurements. 2P-fluxes were estimated using only the first and the last concentration measurement. No outliers were excluded, nor any other quality checks applied to the resulting fluxes. The difference between fluxes was calculated as MP-flux minus 2P-flux.

To test the sample allocation effects on the level of annual balances, we selected sites of high emission, because sites of low emission also show low variability making comparisons arbitrary. For CH_4, a subset that included only closed-chamber measurements from the rewetted fen was chosen for high emission rates. For N_2O we tested a subset that only included measurements from the cropland, where manure applications increased the probability for significant emissions. For both analyses, we presumed that one has resources for a given number of total concentration measurements for each measurement day and that one can either spend them on more spatial replicates of 2P-fluxes or less spatial replicates of MP-fluxes.

The CH_4 data set features six spatial replicates in each of three dominant plant stands (*Phragmites*, *Typha*, *Carex*), resulting in 18 total spatial replications. On these spots, closed-chamber measurements were conducted every two weeks over two years. The N_2O data set features three spatial replicates in each of six agricultural treatment combinations (“T01” to “T06”), again resulting in 18 total spatial replications. On these spots, closed-chamber measurements were conducted every one to two weeks over two years (Year 1: 27–33 measurements per spot; Year 2: 32–47 measurements per spot). Four concentration samples were taken during each chamber closure (in contrast to five in the CH_4 data set).

For CH_4, we first selected measurements separately for each vegetation stand, thus, assuming sampling in rather homogenous ecosystems or systems with previously known influencing factors (*e.g*., factorial designs of agricultural experiments). In this case, we supposed that the total number of concentration measurements for each day and vegetation stand was fixed to 10. These could either be allocated to two 5P-measurements or five 2P-measurements. Thus, we first calculated annual balances (g m^{-2} y^{-1}) for each measurement spot by simple linear interpolation. Then, we randomly picked two out of six replicates for each vegetation stand where we used 5P-fluxes for annual balancing, opposed to picking five (which incorporated the first two) out of six replicates, where we used 2P-fluxes. We repeated this random selection 100 times and calculated the mean annual balance by averaging the results of the individual spots for each selection. As a second test, we tested the performance of 2P- *vs*. MP-fluxes in a system with a larger spatial heterogeneity by randomly sampling across vegetation types. Thus, we now picked two spots from the whole data set (again two spots with 5P-fluxes vs. five spots with 2P-fluxes). Again, we repeated this random selection 100 times and averaged the results for each selection.

We slightly adapted the selection of measurements to the changed study design of the N_2O tests. Again, we first selected measurements separately for each agricultural treatment. Now, we assumed that the number of concentration measurements for each day and treatment was fixed to 4. These could be allocated either to two 2P-measurements or one 4P-measurement. We randomly selected this number of spots once separately for each treatment and once from the whole data set.

Appendix 2: Tool for flux measurement planning

When planning GHG measurements, we propose the following simple approach to determine the optimal balance between spatial coverage and flux precision given the individual amount of available resources: The coefficient of determination (*i.e.*, R^2) is commonly used to describe the goodness-of-fit of the linear regression for each closed-chamber measurement (R^2_{flux}). Thus, R^2_{flux} indirectly includes information on the amount of scatter around the estimated linear regression that would result in a higher uncertainty when reducing the number of concentration measurements. Similarly, the coefficient of determination could be used to describe the variation that results from spatial heterogeneity. In this case, spatial distance forms the independent variable, while gas flux is the dependent variable. Given that there is no spatial gradient in the studied system, a homogenous system would lead to large $R^2_{spatial}$ values and a more heterogeneous system to small values. This is a result of the observed fluxes deviating less or more from the true average flux.

R^2_{flux} and $R^2_{spatial}$ can be used to determine the investment factors that described the amount of effort that should be assigned towards better fluxes (i_{flux}) or better spatial coverage ($i_{spatial}$), where $i_{flux} = 1 - R^2_{flux}$ and $i_{spatial} = 1 - R^2_{spatial}$. Then, the final balance between the two investments can be determined *via* the relation:

$$Balanced\ effort = \frac{i_{flux}}{i_{flux}+i_{spatial}} : \frac{i_{spatial}}{i_{flux}+i_{spatial}}. \quad (1)$$

For example, a relation of 6:4 means that 60% of the available resources should be allocated towards obtaining better gas flux estimates. Since both the number of concentration measurements and the number of chamber placements must be integers (*e.g.*, there is no half concentration measurement), it is possible that the best result involves using differing numbers of concentration measurements per chamber placement. However, the decision whether this is acceptable and/or necessary lies with the researcher. In general, the proposed approach should only be applied when the underlying assumption of linearity is not violated for the modeled relationships.

Table A1: Closed-chamber measurements used for the analyses.

Ecosystem	Time period	Number of measurements	Published in
Cropland	Oct 2012– Sep 2014	*n* = 1669	*Fiedler* et al. (2016): Tillage-induced short-term soil organic matter turnover and respiration. *Soil* 2, 475–486. *Fiedler* et al. (2015): Soil respiration after tillage under different fertiliser treatments – implications for modelling and balancing. *Soil Till. Res*. 150, 30–42.
Peat grassland	Apr 2010– Dec 2011	*n* = 589	*Tiemeyer* et al. (2016): High emissions of greenhouse gases from grasslands on peat and other organic soils. *Global Change Biology* 22, 4134–4149.
Restored bog	Apr 2010– Dec 2011	*n* = 705	*Tiemeyer* et al. (2016): High emissions of greenhouse gases from grasslands on peat and other organic soils. *Global Change Biology* 22, 4134–4149.
Restored fen	Mar 2011– Feb 2013	*n* = 971	*Günther* et al. (2015): The effect of biomass harvesting on greenhouse gas emissions from a rewetted temperate fen. *GCB Bioenergy* 7, 1092–1106. *Günther* et al. (2014): Scale-dependent temporal variation in determining the methane balance of a temperate fen. *Greenhouse Gas Measur. Manage*. 4, 41–48. *Huth* et al. (2013): The effect of an extraordinarily wet summer on methane emissions from a 15-year re-wetted fen in north-east Germany. *Mires Peat* 13, Article 02, 1–7.
Total		*n = 3934*	

Table A2: Mean annual methane balances ± standard deviation (g m^{-2} y^{-1}) of 100 random selections of two MP-fluxes and five 2P-fluxes per measurement day in each vegetation stand ("*Carex*", "*Phragmites*", "*Typha*") and from the whole ecosystem ("Total ecosystem") for the two measurement years. For comparison, values are also given for annual balances of all measurement spots using only MP-fluxes ("All MP-fluxes", grey background).

		MP-fluxes	2P-fluxes	All MP-fluxes
Year 1	*Carex*	54.5 ± 28.4	51.7 ± 42.3	53.0 ± 31.2
	Phragmites	17.1 ± 2.5	18.1 ± 3.3	17.2 ± 2.8
	Typha	17.4 ± 5.6	20.7 ± 9.5	17.3 ± 5.9
	Total ecosystem	28.4 ± 23.8	30.2 ± 27.9	29.1 ± 24.6
Year 2	*Carex*	3.2 ± 1.1	4.4 ± 1.8	3.2 ± 1.3
	Phragmites	1.6 ± 0.6	1.3 ± 0.5	1.6 ± 0.7
	Typha	3.9 ± 1.2	3.3 ± 0.9	3.8 ± 1.4
	Total ecosystem	3.2 ± 1.4	3.0 ± 1.7	2.9 ± 1.4

Table A3: Mean annual nitrous oxide balances ± standard deviation (g m^{-2} y^{-1}) of 100 random s elections of one MP-flux and two 2P-fluxes per measurement day and treatment ("T01" to "T06") and from the whole ecosystem ("Total ecosystem") for the two measurement years. For comparison, mean and standard deviation values are also given for annual balances of all measurement spots using only MP-fluxes ("All MP-fluxes", grey background).

	Treatment	MP-fluxes	2P-fluxes	All MP-fluxes
Year 1	T01	–0.03 ± 0.14	–0.03 ± 0.16	–0.03 ± 0.17
	T02	–0.01 ± 0.12	–0.02 ± 0.13	–0.01 ± 0.15
	T03	0.07 ± 0.08	0.03 ± 0.12	0.06 ± 0.1
	T04	0.02 ± 0.11	0.00 ± 0.13	0.02 ± 0.14
	T05	0.17 ± 0.12	0.12 ± 0.12	0.16 ± 0.15
	T06	0.71 ± 0.1	0.67 ± 0.11	0.69 ± 0.13
	Total ecosystem	0.18 ± 0.3	0.13 ± 0.29	0.15 ± 0.28
Year 2	T01	0.02 ± 0.05	–0.01 ± 0.06	0.01 ± 0.06
	T02	0.12 ± 0.05	0.11 ± 0.05	0.12 ± 0.06
	T03	0.29 ± 0.14	0.26 ± 0.14	0.28 ± 0.17
	T04	0.01 ± 0.04	0.01 ± 0.04	0.01 ± 0.05
	T05	0.11 ± 0.06	0.06 ± 0.01	0.11 ± 0.07
	T06	0.10 ± 0.12	0.13 ± 0.06	0.09 ± 0.15
	Total ecosystem	0.11 ± 0.13	0.10 ± 0.12	0.10 ± 0.13